A3273-8

AUGUST 1990

ACADEMIC MEDICINE

JOURNAL OF THE ASSOCIATION OF AMERICAN MEDICAL COLLEGES

Medical Education in China

Management of Academic Medicine

Development of Medical Education

Medical Library Education

National Examination Validity

Performance on the ABIM Examination

Quarterly Profile Examination

College Health Services as Training Sites

Students' Research and Later Activities

Physicians' Patient Load and Resource Use

Consensus on Nutrition Teaching

National Policy Perspectives

VOLUME 65 • NUMBER EIGHT

AUGUST 1990
VOLUME 65
NUMBER EIGHT

ACADEMIC MEDICINE
JOURNAL OF THE ASSOCIATION OF AMERICAN MEDICAL COLLEGES

Special Articles:	China's Quest for Management of Academic Medicine, *by William T. Butler, M.D., and Steven J. Ruma, Ph.D.*	497
	Medical Library Education in China, *by David S. Crawford, Dip.Lib.Stud., and Dizhi Xiong, M.D.*	502
	Predictive Validity of a National Examination for Medical Graduates in the People's Republic of China, *by Liu Bingxun, M.D., Mei Renlang, M.D., Niu Weidong, M.D., Zhang Zu, M.D., Qui Xiangxing, M.D., and Wang Lie, M.D.*	505
International Medical Education	Development of Medical Education in China, *by Deng Yizhong, M.D.*	512
Ideas for Medical Education	The Quarterly Profile Examination, *by Louise Arnold, Ph.D., and T. Lee Willoughby, M.S.*	515
From the President		517
National Policy Perspectives		518
AAMC Focus		520
Commentary	Medical Research and Research!America, *by Lowell Weicker, Jr.*	521
Letter to the Editor	Revitalizing the Undergraduate College–Medical School Linkage, *by William A. Thomson, Ph.D., Leslie M. Miller, Ph.D., Quentin W. Smith, M.S., and Zenaido Camacho, Ph.D.*	522
Research Reports **ARTICLES**	Performances of U.S. Osteopathic and Canadian Medical School Graduates on the American Board of Internal Medicine Certifying Examinations, 1984–1988, *by Judy A. Shea, Ph.D., John J. Norcini, Ph.D., and John A. Benson, Jr., M.D.*	523

College Health Services as a Site for Ambulatory Care Training of Residents and Medical Students, *by John M. Chuck, M.D., Robert C. Davidson, M.D., M.P.H., and Suzanne Snively, M.D.* 527

The Association between Students' Research Involvement in Medical School and Their Postgraduate Medical Activities, *by Scott Segal, M.D., Tom Lloyd, Ph.D., Peter S. Houts, Ph.D., Paula L. Stillman, M.D., Robert L. Jungas, Ph.D., and Robert B. Greer III, M.D.* 530

Physicians' Patient Load per DRG, the Consumption of Hospital Resources, and the Incentives of the DRG Prospective Payment System, *by Eric Munoz, M.D., M.B.A., Katherine Mulloy, Jonathan Goldstein, M.P.A., Jordon Josephson, M.D., Noel Tenenbaum, M.D. and Leslie Wise, M.D.* 533

SHORT PAPERS

Priorities for Nutrition Content within a Medical School Curriculum: A National Consensus of Medical Educators, *by the American Society for Clinical Nutrition Committee on Medical/Dental School and Residency Nutrition Education* 538

A Survey of the Personal Computer Laboratory Facilities and Policies at U.S. Medical Schools, *by Linda K. Gunzburger, Ph.D., and Thomas Loesch, M.Ed.* 540

Brief Communications

Amount and Distribution of Interns' Working Hours, *by Frank J. Kroboth, M.D., Alan R. Whitman, Michael Karpf, M.D., and Gerald S. Levey, M.D.* 542

Two Methods of Including Postgraduate Students as Observers of Psychotherapy, *by Hernan Davanzo, M.D., Claudia Pruzzo, M. Teresa Sanfuentes, Gabriela Trivelli, and Verónica Wigdorsky* 542

The Effect of Surgery Clerkship Timing on Fourth-Year Students' Surgery Knowledge, *by Frank A. Baciewicz, Jr., M.D., Judy Fagley, Michael Weaver, M.S.N., Richard Yeasting, Ph.D., and Neil R. Thomford, M.D.* 543

Erratum 543

Publishing Services: Hanley & Belfus, Inc.
210 South 13th Street, Philadelphia, PA 19107

AUGUST 1990
VOLUME 65
NUMBER EIGHT

ACADEMIC MEDICINE

JOURNAL OF THE ASSOCIATION OF AMERICAN MEDICAL COLLEGES

EDITOR: Addeane S. Caelleigh

DEPUTY EDITOR: Albert Bradford

SUBMISSIONS EDITOR: Christopher Greene

EDITORIAL ASSISTANT: Lisa R. Dittrich

PUBLISHER'S EDITOR: Sandra A. Lovegrove

ASSOCIATE EDITORS
Senior Associate Editor
Donald G. Kassebaum, M.D.
Edward J. Stemmler, M.D.
Association of American Medical Colleges

Ideas in Medical Education
Janine C. Edwards, Ph.D.
St. Louis University, School of Medicine

International Medical Education
David S. Greer, M.D.
Brown University Program in Medicine

Informatics
Daniel R. Masys, Ph.D.
Lister Hill National Center for Biomedical Communications

Book Reviews
Sherman Mellinkoff, M.D.
University of California, Los Angeles School of Medicine

Academic Computing in Medicine
Robert F. Rubeck, Ph.D.
University of Kentucky College of Medicine

EDITORIAL BOARD
Chairman
Milton Corn, M.D.
Georgetown University Hospital
Philip Anderson, M.D.
University of Missouri-Columbia School of Medicine
Nancy Bennett, Ph.D.
Harvard Medical School
Eta Berner, Ed.D.
University of Alabama School of Medicine
Ruth Bulger, Ph.D.
National Academy of Sciences
Vincent A. Fulginiti, M.D.
Tulane University School of Medicine
Nancy Gary, M.D.
Albany Medical College
Beatrix Hamburg, M.D.
Mt. Sinai Hospital
Murray M. Kappelman, M.D.
The University of Maryland School of Medicine
Ralph W. Muller
University of Chicago Hospitals
Herbert Pardes, M.D.
College of Physicians and Surgeons of Columbia University
James A. Pittman, Jr., M.D.
University of Alabama School of Medicine
W. Vickery Stoughton
Toronto General Hospital
Martha Vaughan, M.D.
National Institutes of Health

OFFICERS OF THE ASSOCIATION
Chairman
David H. Cohen, Ph.D.
Northwestern University

Chairman-Elect
William T. Butler, M.D.
Baylor College of Medicine

Immediate Past Chairman
D. Kay Clawson, M.D.
University of Kansas Medical Center School of Medicine

President
Robert G. Petersdorf, M.D.

Chairman, Council of Deans
L. Thompson Bowles, M.D., Ph.D.
George Washington University School of Medicine and Health Sciences

Chairman, Council of Academic Societies
Joe Dan Coulter, Ph.D.
University of Iowa College of Medicine

Chairman, Council of Teaching Hospitals
Raymond G. Schultze, M.D.
University of California, Los Angeles Medical Center

Academic Medicine (ISSN 1040-2446) is owned and published monthly by the Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036. Publishing services by Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, Pennsylvania 19107. Second-class postage paid at Washington, DC and additional mailing offices.

Copyright 1990 by the Association of American Medical Colleges. All material subject to this copyright may be photocopied for the noncommercial purpose of scientific or educational advancement. Printed in the U.S.A.

Subscription rates: U.S., Canada, Latin America—$60.00 per year and $7.00 per single copy ($30.00 per year for students); all other locations—$70.00 per year and $8.00 per single copy.

All advertising material published in this journal is intended to conform to ethical standards, but acceptance does not imply endorsement. The views expressed in all articles and features in this journal are those of the authors and do not necessarily reflect the position of the AAMC.

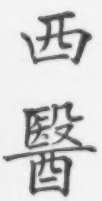
CORRESPONDENCE
Manuscripts: The Editor, *Academic Medicine*, Association of American Medical Colleges, One Dupont Circle, NW, Washington, DC 20036.

Advertising: Diane Sherel, Advertising Manager, Hanley & Belfus, Inc., 210 South 13th Street, Philadelphia, PA 19107; (215) 546-0313.
Subscriptions and changes of address: Membership and Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036; (202) 828-0416.
Cancellations of an *Academic Medicine* subscription must be submitted in writing to Publication Orders, AAMC, One Dupont Circle, NW, Washington, DC 20036. Claims for back issues will be honored for only the three months before the cancellation.
Reprints: Request reprints directly from the authors of articles, not from *Academic Medicine*.
POSTMASTER: Send address changes to *Academic Medicine*, AAMC, Suite 200, One Dupont Circle, NW, Washington, DC 20036.

LIU BINGXUN, M.D., MEI RENLANG, M.D., NIU WEIDONG, M.D.,
ZHANG ZU, M.D., QIU XIANGXING, M.D., and WANG LIE, M.D.

Predictive Validity of a National Examination for Medical Graduates in the People's Republic of China

Abstract—National examinations for medical graduates were introduced on an experimental basis in the People's Republic of China in 1982. To estimate the predictive validity of the National Medical Examination (NME), an investigation of the postgraduate competence of a sample of the participating examinees was conducted in 1984. The sample consisted of 1,717 of the 4,995 graduates from 13 medical colleges who had taken the initial NME. Their scores on the NME and the ratings given them by directors of postgraduate programs in nine aspects of clinical competence were compared by frequency distribution and product-moment correlation coefficients. Scores on the NME were consistent with measures of postgraduate clinical competence and, as a whole, correlated significantly with the ratings of clinical competence, supporting the use of the score on the NME as a predictor of postgraduate clinical competence. However, the extent of the relationship between the NME score and postgraduate clinical competence varied according to the specialty program of postgraduate medical training. *Acad. Med.* 65(1990):505–511.



To accumulate the experience needed to establish a system of national medical qualifying examinations and to evaluate objectively the educational quality of the medical schools in the People's Republic of China, the Ministry of Public Health has administered National Medical Examinations (NMEs) to graduates of some of these medical schools annually since 1982.

The NMEs have each been composed of 200 to 300 multiple-choice questions about internal medicine, surgery, obstetrics-gynecology, and pediatrics, to be completed in six hours. Test items are selected from the Test Item Bank developed by the Committee of Test Item Formulation for the National Medical Examination.

When the first NME was given, in 1982, 4,995 medical graduates from 13 medical colleges were asked to take this examination. Two hundred test items were put to the examinees; their average score was 169.5. The initial results of this examination were analyzed and found to be content-valid and reliable, with a Kuder-Richardson (KR-20) reliability coefficient of 0.78.[1] A second study, of 226 examinees, demonstrated that the product-moment correlations between scores on the NME and undergraduate performances in basic and clinical sciences in the medical school were 0.48 and 0.43, respectively.[2] Two other studies demonstrated reliability coefficients of 0.58 and 0.65, and correlations with undergraduate performances in medical schools ranging from 0.26 to 0.69.[3,4]

Whether the results of such an examination can be used as a predictor of the postgraduate clinical competence of the examinee has been debated. Kappelman[5] argued that multiple-choice questions could be used to measure examinees' mastery of knowledge, but could not be used to assess such clinical abilities as appropriate interaction with patients, skill in accessing the literature, and technical skills. Various investigators[6–8] have studied the predictive validity of qualifying examinations. Some concluded that an important relationship existed between the scores on a nationally standardized examination and ratings of the postgraduate clinical competence of the examinees, while others reached an utterly contrary conclusion. Gonnella and Veloski and their colleagues[6,7] reported a positive association between scores on the examinations of the National Board of Medical Examiners (NBME) and ratings of clinical competence at the end of the first postgraduate year; however, the predictive validity of the NBME examination scores depended upon the types of specialty training programs that the examinees were taking.[6,7] In order to ascertain the predictive validity of the first NME, we investigated the postgraduate clinical competence of a sample of the examinees in the fall of 1984.

Method

The sample for the present study consisted of 1,717 residents who graduated from 13 medical schools and took the first NME in 1982. Of the total group, 799 were training in internal medicine, 499 in surgery, 260 in obstetrics-gynecology, and 159 in pediatrics. They constituted 34.4% of the 4,995 examinees taking the NME in that year. The other 3,278 medical

Dr. Liu Bingxun is director, Department of Medical Education, and chairman of the Committee on Test Item Formulation for National Medical Examination, Ministry of Public Health, Beijing, People's Republic of China; Dr. Mei Renlang is associate professor of medical education, Shanghai Medical University; Dr. Niu Weidong is director, Division of Research and Statistics, National Medical Examination Center; Dr. Zhang Zu is research assistant, Division of Medical Education Research, Norman Bethune Medical University; Dr. Qiu Xiangxing is director, Division of Medical Education, Shanghai Second Medical University; and Dr. Wang Lie is research assistant, Medical Education Research Center, China Medical University.

Correspondence should be addressed to Dr. Liu, Director, Department of Medical Education, Ministry of Public Health, People's Republic of China, 44 Houhai Beiyan, Beijing, China. Reprints will not be available.

The message of the calligraphy is "Chinese traditional medicine."

graduates were taking postgraduate training in basic medical sciences or other clinical specialty programs, or were engaged in other activities, and were therefore excluded from the study.

Rating forms (one for each resident) were sent to the residency supervisors, who were asked to rate the clinical competences of their trainees on a four-point scale ranging from 1 for unsatisfactory to 4 for excellent. The form included 44 items that described nine aspects of clinical competence: data gathering, physical examination, differential diagnosis, patient management, clinical procedures, application of basic science knowledge, verbal communication, service attitude, and comprehensive assessment. ("Comprehensive assessment" was not the sum or the average value of the ratings of the other eight aspects, but a general evaluation of a resident based on his or her clinical performance as a whole and the general impression on the rater.) The average value of the ratings of all items on each aspect of competence was computed for each resident; the residents were then classified accordingly into three ability groups: those whose average ratings on a particular aspect were in the top 15% were put into the high-ability group for that aspect, while those in the lowest 15% were considered to be in the low-ability group. The remaining residents constituted the medium-ability group.

In the same way, according to their scores on the NME, the residents were again divided into three groups: high-score group, medium-score group, and low-score group. The two sets of data were compared statistically by their joint frequency distributions and Pearson product-moment correlation coefficients.

Results

Table 1 shows (1) the numbers of graduates classified by their NME scores (low, medium, or high) and their postgraduate clinical competence ratings (low-, medium-, or high-ability) for each of the nine aspects of clinical competence (the count); (2) the percentage distributions of the postgraduate competence ratings, by the NME scores (the column percentage); (3) the percentage distributions of the NME scores, by postgraduate competence ratings (the row percentage); and (4) chi-square (χ^2) for each small data matrix.

Column Percentages

An analysis of the percentage distributions of the levels of ratings of postgraduate clinical competence

Table 1: Numbers of Graduates and Percentage Distributions of National Medical

Aspect of Competence	Rating	No. of Graduates (n = 1,717)		Total			
				High (n = 271)	Med. (n = 1,163)	Low (n = 283)	χ^2
Comprehensive assessment	High	261	Count	50	192	19	42.322*
			Col Pct	18	17	7	
			Row Pct	**19**	**74**	**7**	
	Med.	1,211	Count	199	815	197	
			Col Pct	73	70	70	
			Row Pct	**16**	**67**	**16**	
	Low	245	Count	22	156	67	
			Col Pct	8	13	24	
			Row Pct	**9**	**64**	**27**	
Data gathering	High	828	Count	144	585	99	29.902*
			Col Pct	53	50	35	
			Row Pct	**17**	**71**	**12**	
	Med.	579	Count	88	383	108	
			Col Pct	32	33	38	
			Row Pct	**15**	**66**	**19**	
	Low	310	Count	39	195	76	
			Col Pct	14	17	27	
			Row Pct	**13**	**63**	**25**	
Physical examination	High	639	Count	113	452	74	37.495*
			Col Pct	42	39	26	
			Row Pct	**18**	**71**	**12**	
	Med.	489	Count	82	338	69	
			Col Pct	30	29	24	
			Row Pct	**17**	**69**	**14**	
	Low	589	Count	76	373	140	
			Col Pct	28	32	49	
			Row Pct	**13**	**63**	**24**	

within different levels of NME scores demonstrated that the two were consistent with each other in most aspects; the exceptions were data gathering and differential diagnosis (see the column percentage distribution in Table 1).

The proportion of graduates with high NME scores was comparatively greater in the groups with high clinical ability than in the groups with low clinical ability. In addition, the proportion of graduates with low NME scores was comparatively larger in the groups with low clinical competency ratings than in those with high ratings. So with regard to comprehensive assessment of clinical competence, 18% of the residents in the group with high total scores were identified as members of the high-ability group, but only 8% were rated as having low ability. Conversely, only 7% of the residents in the group with low total scores were classified into the group with high ability, while up to 24% were rated as having low ability.

However, with respect to data-gathering skills, the proportion of graduates with low NME scores was comparatively larger in the group with high ability than in the group with low ability. Thirty-five percent of the residents in the low-total-score group were considered to have high ability, while only 27% of them were put into the low-ability group. Conversely, only 26% of the residents with high total scores were in the high-ability group for differential diagnosis, while the corresponding percentage of those with high total scores and in the low-ability group was as high as 40%. The same phenomenon was observed in the relationships between the ratings on these two competence aspects and the NME scores in various disciplines.

Row Percentages

Table 1 also displays the relationships among the variables of postgraduate competence and NME scores in another way—by row percentage distribution, that is, the percentage distribution of NME scores by postgraduate competence ratings. The ratings of clinical competence after graduation tended to be consistent with the NME scores in all the nine aspects, including data gathering and differential diagnosis. For example, in the comprehensive assessment area, 19% of the graduates in the high-ability group achieved high total scores on the NME, but only 7% of them belonged to the group with low total NME scores. At the same time, only 9% of the graduates with low ability ratings achieved good performances on the NME, while the percentage in the group who had low NME

Examination Scores and Postgraduate Competence Ratings, People's Republic of China, 1984

National Medical Examination Score															
Int. Med. Component				Surgery Component				Ob.-Gyn. Component				Pediatrics Component			
High (n = 347)	Med. (n = 1,058)	Low (n = 312)	χ^2	High (n = 307)	Med. (n = 1,098)	Low (n = 312)	χ^2	High (n = 267)	Med. (n = 1,140)	Low (n = 310)	χ^2	High (n = 488)	Med. (n = 880)	Low (n = 349)	χ^2
59	169	33	25.210*	61	172	28	26.733*	48	172	41	5.080	82	146	33	17.538*
17	16	11		20	16	9		18	15	13		17	17	9	
23	**65**	**13**		**23**	**66**	**11**		**18**	**66**	**16**		**31**	**56**	**13**	
254	747	210		214	779	218		184	812	215		347	615	249	
73	71	67		70	71	70		69	71	69		71	70	71	
21	**62**	**17**		**18**	**64**	**18**		**15**	**67**	**18**		**29**	**51**	**21**	
34	142	69		32	147	66		35	156	54		59	119	67	
10	13	22		10	13	21		13	14	17		12	14	19	
14	**58**	**28**		**13**	**60**	**27**		**14**	**64**	**22**		**24**	**49**	**27**	
187	509	132	14.075*	139	535	124	17.105*	150	549	129	13.041†	252	431	145	11.729*
54	48	42		55	49	40		56	48	42		52	49	42	
23	**61**	**16**		**20**	**65**	**15**		**18**	**66**	**16**		**30**	**52**	**18**	
110	364	105		88	376	115		74	391	114		156	300	123	
32	34	34		29	34	37		28	34	37		32	34	35	
19	**63**	**18**		**15**	**65**	**20**		**13**	**68**	**20**		**27**	**52**	**21**	
50	185	75		50	187	73		43	200	67		80	149	81	
14	17	24		16	17	23		16	18	22		16	17	23	
16	**60**	**24**		**16**	**60**	**24**		**14**	**65**	**22**		**26**	**48**	**26**	
135	408	96	15.890*	126	424	89	26.516*	113	424	102	15.933*	210	319	110	19.335*
39	39	31		41	39	29		42	37	33		43	36	32	
21	**64**	**15**		**20**	**66**	**14**		**18**	**66**	**16**		**33**	**50**	**17**	
109	299	81		89	322	78		88	317	84		134	264	91	
31	28	26		29	29	25		33	28	27		27	30	26	
22	**61**	**17**		**18**	**66**	**16**		**18**	**65**	**17**		**27**	**54**	**19**	
103	351	135		92	352	145		66	399	124		144	297	148	
30	33	43		30	32	46		25	35	40		39	34	42	
17	**60**	**23**		**16**	**60**	**25**		**11**	**68**	**21**		**24**	**50**	**25**	

Continued—

Table 1: Numbers of Graduates and Percentage Distributions of National Medical

Aspect of Competence	Rating	No. of Graduates (n = 1,717)		Total High (n = 271)	Total Med. (n = 1,163)	Total Low (n = 283)	χ^2
Differential diagnosis	High	434	Count	72	317	45	20.341*
			Col Pct	26	27	16	
			Row Pct	**17**	**73**	**10**	
	Med.	528	Count	91	361	86	
			Col Pct	34	30	30	
			Row Pct	**17**	**66**	**16**	
	Low	755	Count	108	495	152	
			Col Pct	40	43	54	
			Row Pct	**14**	**66**	**20**	
Patient management	High	318	Count	60	231	27	39.251*
			Col Pct	22	20	10	
			Row Pct	**19**	**73**	**8**	
	Med.	1,159	Count	184	786	189	
			Col Pct	68	68	67	
			Row Pct	**16**	**68**	**16**	
	Low	240	Count	27	146	67	
			Col Pct	10	13	24	
			Row Pct	**11**	**61**	**28**	
Clinical procedures	High	315	Count	58	225	32	25.286*
			Col Pct	21	19	11	
			Row Pct	**20**	**71**	**9**	
	Med.	1,119	Count	176	763	180	
			Col Pct	65	66	63	
			Row Pct	**17**	**68**	**15**	
	Low	283	Count	37	175	71	
			Col Pct	14	15	25	
			Row Pct	**15**	**62**	**23**	
Application of basic science knowledge	High	316	Count	68	222	26	43.255*
			Col Pct	25	19	9	
			Row Pct	**22**	**70**	**8**	
	Med.	1,130	Count	178	766	186	
			Col Pct	66	66	66	
			Row Pct	**16**	**68**	**16**	
	Low	271	Count	25	175	71	
			Col Pct	9	15	25	
			Row Pct	**9**	**65**	**26**	
Verbal communication	High	460	Count	76	327	57	28.986*
			Col Pct	28	28	20	
			Row Pct	**17**	**71**	**12**	
	Med.	1,004	Count	169	677	158	
			Col Pct	62	58	56	
			Row Pct	**17**	**67**	**16**	
	Low	253	Count	26	159	68	
			Col Pct	10	14	24	
			Row Pct	**10**	**63**	**27**	
Service attitude	High	597	Count	112	408	77	25.785*
			Col Pct	41	35	27	
			Row Pct	**19**	**68**	**13**	
	Med.	708	Count	110	489	109	
			Col Pct	41	42	39	
			Row Pct	**16**	**69**	**15**	
	Low	412	Count	49	266	97	
			Col Pct	18	23	34	
			Row Pct	**12**	**65**	**24**	

*p = .01 for values of 13.277 or higher.
†p = .05 for values of 9.488 or higher.

Examination Scores and Postgraduate Competence Ratings, People's Republic of China, 1984 — *Continued*

National Medical Examination Score															
Int. Med. Component				Surgery Component				Ob.-Gyn. Component				Pediatrics Component			
High (n = 347)	Med. (n = 1,058)	Low (n = 312)	χ^2	High (n = 307)	Med. (n = 1,098)	Low (n = 312)	χ^2	High (n = 267)	Med. (n = 1,140)	Low (n = 310)	χ^2	High (n = 488)	Med. (n = 880)	Low (n = 349)	χ^2
88	286	60	12.360†	84	290	60	13.540*	73	285	76	7.210	134	227	73	7.478
25	27	19		27	26	19		27	25	25		27	26	21	
20	**66**	**14**		**19**	**67**	**14**		**17**	**66**	**18**		**31**	**52**	**17**	
121	311	96		95	346	87		95	346	87		158	263	107	
35	29	31		31	32	28		35	30	28		32	30	31	
23	**59**	**18**		**18**	**66**	**16**		**18**	**66**	**16**		**30**	**50**	**20**	
138	461	156		128	462	165		99	509	147		196	390	169	
40	44	50		42	42	53		37	45	47		40	44	48	
18	**61**	**21**		**17**	**61**	**22**		**13**	**68**	**19**		**26**	**52**	**22**	
72	205	41	27.190*	74	206	38	22.305*	60	212	46	9.075	103	164	51	17.312*
21	19	13		24	19	12		22	19	15		21	19	15	
23	**64**	**13**		**23**	**65**	**12**		**19**	**67**	**14**		**32**	**52**	**16**	
244	712	203		196	751	212		178	771	210		335	594	230	
70	67	65		64	68	68		67	68	68		69	68	66	
21	**61**	**18**		**17**	**65**	**18**		**15**	**67**	**18**		**29**	**51**	**20**	
31	141	68		37	141	62		29	157	54		50	122	68	
9	13	22		12	13	20		11	14	17		10	14	19	
13	**59**	**28**		**15**	**59**	**26**		**12**	**65**	**23**		**21**	**51**	**28**	
67	205	43	8.951	72	208	35	25.064*	62	201	52	15.510*	100	160	55	9.605†
19	19	14		22	19	11		23	17	17		20	18	16	
22	**65**	**13**		**23**	**66**	**10**		**22**	**62**	**16**		**32**	**51**	**17**	
229	686	204		199	714	206		174	756	189		327	562	230	
66	65	65		65	65	66		65	66	61		67	64	66	
21	**62**	**17**		**19**	**64**	**17**		**17**	**67**	**16**		**30**	**50**	**20**	
51	167	65		36	176	71		31	183	69		61	158	64	
15	16	21		12	16	23		12	16	22		13	18	18	
18	**60**	**21**		**15**	**62**	**23**		**13**	**65**	**23**		**23**	**57**	**21**	
70	208	38	24.279*	71	210	35	31.377*	63	208	45	8.865	112	160	44	25.395*
20	20	12		23	19	11		24	18	15		23	18	13	
22	**66**	**12**		**22**	**66**	**11**		**20**	**66**	**14**		**35**	**51**	**14**	
240	687	203		202	726	202		168	753	209		322	577	231	
69	65	65		66	66	65		63	66	67		66	66	66	
21	**61**	**18**		**18**	**64**	**18**		**15**	**67**	**18**		**28**	**51**	**20**	
37	163	71		34	162	75		36	179	56		54	143	74	
11	15	23		11	15	24		13	16	18		11	16	21	
14	**60**	**26**		**13**	**60**	**28**		**13**	**66**	**21**		**20**	**53**	**27**	
95	297	68	23.683*	89	299	72	14.275*	86	295	79	9.154	131	254	75	15.876*
27	28	22		29	27	23		32	26	25		27	29	21	
21	**65**	**15**		**19**	**65**	**16**		**19**	**64**	**17**		**28**	**55**	**16**	
216	615	173		188	639	177		154	673	177		288	514	202	
62	58	56		61	58	57		58	59	57		59	58	58	
22	**61**	**17**		**19**	**64**	**18**		**15**	**67**	**18**		**29**	**51**	**20**	
36	146	71		30	160	63		27	172	54		69	112	72	
10	14	23		10	15	20		10	15	17		14	13	21	
14	**58**	**28**		**12**	**63**	**25**		**11**	**68**	**21**		**27**	**44**	**28**	
138	364	95	14.876*	114	395	88	15.512*	100	394	103	8.254	189	309	99	17.447*
40	34	30		37	36	28		37	35	33		39	35	28	
23	**61**	**16**		**19**	**66**	**15**		**17**	**66**	**17**		**32**	**52**	**17**	
142	446	120		128	456	124		115	477	116		196	372	140	
41	42	38		42	42	40		43	42	38		40	42	40	
20	**63**	**17**		**18**	**64**	**18**		**16**	**67**	**17**		**28**	**53**	**20**	
67	248	97		65	247	100		52	269	91		103	199	110	
19	23	31		21	22	32		19	24	29		21	23	32	
16	*60*	*24*		*16*	*60*	*24*		*13*	*65*	*22*		*25*	*48*	*27*	

total scores was as high as 27%. The relationships between the scores in the various disciplines covered in the NME and the postgraduate clinical competence ratings were in the same direction.

NME Predictive Validity

In order to identify the predictive validity of the NME, the Pearson correlation coefficients between the scores on the NME and the ratings in each competence aspect were calculated. It was found that the NME total scores correlated with the ratings of clinical competence in all the nine aspects. The coefficients ranged from 0.1205 to 0.1946 and were statistically significant ($p < .01$). Generally, statistically significant correlations were also found between the NME total scores in the various specialty components and the ratings of the various clinical competences, although the correlations among a few variables, in obstetrics-gynecology in particular, were of lesser magnitude.

In exploring the NME's predictive validity, the individual specialty programs that the residents were pursuing were taken into consideration. Significant correlations were found among the variables for the group of residents training in internal medicine programs, except for the correlation between the NME score on obstetrics-gynecology and competence in physical examination. However, the correlations in the other groups did not follow this pattern. For surgery and obstetrics-gynecology, half of these correlation coefficients were not statistically significant, including those relating to the score of the same discipline in the NME. No significant correlation was found among the variables in the pediatrics group except that between the NME score in pediatrics and the rating of competence in making a diagnosis.

Discussion

It is important to clarify the relationship between NME scores and postgraduate clinical competence in assessing the validity of this measuring instrument. In educational measurement, it is desirable that a valid and reliable examination not only determine the current level of performance of the examinees but also have predictive value regarding the development of their professional careers. Since the concurrent validity and reliability of the NME had been confirmed in 1982, the authors hypothesized that those graduates who scored high on the examination should receive high clinical competence ratings and vice versa.

To test the hypothesis, two statistical approaches were used. The data presented in Table 1 provide evidence supporting the hypothesis. In addition, significant product-moment correlations were found among the various aspects of the NME total score and the ratings of postgraduate clinical competence. These findings are consistent with other reports.[7] Therefore, generally speaking, it seems that the NME scores can be useful in predicting postgraduate competence.

As shown in the column percentage distribution in Table 1, the ratings of two aspects of competence—data gathering and differential diagnosis—appear to have been inconsistent with NME scores. "Low scores but high ability" and "high scores but poor ability" were found in these two areas. But these phenomena may have been caused by differences in the criteria used in rating. Raters had lower standards for the aspect of data gathering, so more residents were rated into the high-ability group; this interpretation is supported by negative skewness in the distribution. Conversely, higher standards obtained for differential diagnosis; thus, more residents were rated as having low ability; this is supported by positive skewness in the distribution.

Comparison of the percentage distributions of the examinees in the high-score group and those in the low-score group in these two aspects shows that more examinees in the low-score group were really rated as having high ability than were rated as having low ability in data gathering (35% versus 27%). The former is 1.3 times as large as the latter. But the percentage of examinees rated as having high ability in the high-score group is even greater (53%) and is 3.8 times as large as the percentage for those rated as having low ability (14%) in this aspect.

Similarly, in the ratings of differential diagnosis, more examinees were classified in the group of low ability than in the group of high ability in the high-score group (40% versus 26%), and the former is 1.5 times as large as the latter. But the percentage of examinees rated as having low ability in the low-score group is even larger (54%) and 3.4 times as large as the percentage for those rated as having high ability (16%) in this area.

These false appearances can be clarified by another presentation shown in the row percentage distribution in Table 1, i.e., the percentage distribution of NME scores by postgraduate competence ratings. The table shows that in the data-gathering and differential diagnosis aspects, as in other aspects, the NME scores and competence ratings were in the same direction. There were more examinees with high NME scores in the high-ability groups, and the percentages are 17% versus 12% and 17% versus 10%, respectively. On the contrary, in the groups that had low ability, the proportions of examinees with high scores were smaller than those of examinees with low scores; the percentages are 13% versus 25% and 14% versus 20%.

In analyzing the relationships of the NME specialty component scores to the measures of postgraduate competence in different specialty training programs, we found a significant correlation for internal medicine, while the ratings on half of the clinical competence aspects in surgery and in obstetrics-gynecology did not significantly correlate with the NME scores. No significant correlation was found among almost all the variables in the pediatrics programs. The reasons for these findings may be as follows:

1. They may have occurred only by chance. It seems that chance is unlikely, however, because Gonnella's

study[6] of the relationship between school performance and postgraduate competence resulted in a similar pattern.

2. The sample may not have been large enough. Statistically, the larger the sample, the more likely that the correlation will be significant. In this study, the highest correlation was found among the variables in the internal medicine group, which happened to have the largest sample size compared with the other groups. For surgery, obstetrics-gynecology, and pediatrics, the levels of significance of the correlations among the variables increased with increasing sample size.

3. Different departments and raters may have had different standards even though they had been given a uniform checklist. The Shanghai First Medical College analyzed the consistency of ratings among different medical units and found that there were significant differences among them, which probably explains the variation in results. To improve inter-rater reliability, it seems that more detailed and concrete rating standards should be developed and the number of raters for each resident should be increased.

4. Less variation and a narrower range in grouping the NME scores and the measures of clinical competence create more homogeneity in the residents' groups. When distinctions are not clear enough, especially in a small sample, correlation coefficients will be quite small. If this is the case, other evaluation indexes should be considered in assessing the predictive validity of the NME.

Conclusions

The results of the present study confirm that there is a certain consistency between the NME scores and postgraduate clinical competence, suggesting that the score on the NME can be used as a predictor of postgraduate clinical competence. However, the study has also shown that the extent of the relationship between the NME score and postgraduate clinical competence varies with different specialty programs of postgraduate medical training.

References

*1. The Office of National Medical Examination, Ministry of Public Health, People's Republic of China. *The Report on a Tentative National Examination for Medical Graduates from 13 Medical Schools in 1982.* 1983.

*2. Fang Chunwang, et al. Analyses of the Results of the National Examination for Medical Graduates from Shandong Medical College in 1982. *Med. Educ.* **2**(1983):11-21.

*3. Mei Renlang, et al. *Analysis of the Reliability and Validity of the National Examination.* The Symposium on Medical Education, Shanghai First Medical College, 1983.

*4. Qiu Xiangxing, et al. *The Analysis of the Scores of the Medical Graduates in 1982 National Examination.* The Symposium on Medical Education, Shanghai Second Medical College, 1983.

5. Kappelman, M. M. The Impact of External Examinations on Medical Education Programs and Students. *J. Med. Educ.* **58**(1983):300-308.

6. Gonnella, J. S., and Hojat, M. Relationship Between Performance in Medical School and Postgraduate Competence. *J. Med. Educ.* **58**(1983):679-685.

7. Veloski, J., Herman, M. W., Gonnella, J. S., Zeleznik, C., and Kellow, W. F. Relationships Between Performance in Medical School and First Postgraduate Year. *J. Med. Educ.* **54**(1979):909-916.

8. Hubbard, J. P. *Measuring Medical Education.* 2d ed. Philadelphia, Pennsylvania: Lea and Febiger, 1979.

*Printed in Chinese

Performances of U.S. Osteopathic and Canadian Medical School Graduates on the American Board of Internal Medicine Certifying Examinations, 1984–1988

JUDY A. SHEA, Ph.D., JOHN J. NORCINI, Ph.D., and JOHN A. BENSON, Jr., M.D.

Abstract—The performances of Canadian medical school graduates and U.S. osteopathic medical school graduates who first took the American Board of Internal Medicine Certifying Examination between 1984 and 1988 were compared with the performances, during the same period, of U.S. and foreign medical school graduates. Approximately 100 Canadian graduates took the examination each year; by 1988 the number of osteopathic graduates taking it was 102, double the number participating in 1984. Nearly all the Canadian graduates attended university or university-affiliated residencies, whereas half of the osteopathic graduates attended non-university-affiliated programs. For their overall clinical competence and for the eight components of clinical competence, the Canadian graduates were rated highest, followed by the U.S., osteopathic, and foreign graduates. The Canadians' average examination scores were also highest. The authors discuss the relatively low level of performance of the osteopathic graduates, but conclude that these graduates appear to be an untapped source of talented physicians for internal medicine residencies. The limitations of studying self-selected groups of candidates are also discussed. *Acad. Med.* 65(1990): 523–526.

Several studies have shown that physicians' performances during residency and at the completion of training are related to the types of medical school experiences they have had.[1–3] For example, graduates of medical schools in the United States (USMGs) consistently score higher than do graduates of foreign medical schools (FMGs) on the American Board of Internal Medicine (ABIM) certifying examination.[2–5] The performance differentials between USMGs and FMGs are also observed for evaluations of clinical competence given by program directors at the conclusion of residency training. Breakdowns of data for subgroups of FMGs show finer distinctions in performance. For example, citizens of the United States who attended a medical school in the Caribbean had lower scores and ratings of competence, in general, than did both foreign citizens attending foreign schools and U.S. citizens attending foreign schools in locations other than the Caribbean.[3]

Other studies have shown that performance on the ABIM certifying examination is related to the quality of the residency program, regardless of where one has attended medical school.[6–8] The advantage of attending a residency viewed as outstanding by experts in the field was observed for men and women who were first-time takers of the certifying examination and were from the combined group of United States and Canadian medical school graduates.[9] In addition, foreign medical graduates who had attended a university residency program did better on the certifying examination than did their peers who had attended a non-university-affiliated program.[2]

Overall, past studies provide data that guide expectations about groups of candidates seeking certification in internal medicine. Such information may be helpful to program directors and other professionals involved in the graduate medical education process. In order to more completely describe the relationship between type of undergraduate medical school and later performance, it would be useful to have information about two previously unstudied subgroups of candidates: graduates of Canadian medical schools and graduates of osteopathic medical schools. (All osteopathic medical schools are located in the United States.)

Each year, both Canadian and osteopathic medical schools graduate hundreds of physicians. In 1987, 1,763 physicians completed undergraduate medical training in Canada, and 1,579 received degrees in osteopathic medicine in the United States.[10,11] Available data suggest that sizeable numbers of graduates from both backgrounds take graduate

Dr. Shea is senior psychometrician, Dr. Norcini is executive vice-president for evaluation and research, and Dr. Benson is president at the American Board of Internal Medicine. Drs. Shea and Norcini are in Philadelphia, Pennsylvania, and Dr. Benson is in Portland, Oregon.

Correspondence and requests for reprints should be addressed to Dr. Shea, American Board of Internal Medicine, 3624 Market Street, Philadelphia, PA 19104.

training in internal medicine. As of September 1, 1987, a total of 436 osteopathic graduates were in internal medicine residency programs approved by the Accreditation Council for Graduate Medical Education.[12] In the same year, 1,256 residents were in training in internal medicine in a Canadian program, although the proportion of Canadian graduates is unreported.[10] Further study is needed because these data neither include Canadian graduates training in residency programs in the United States nor allow for the fact that not all trainees attempt voluntary certification.

This study details the performances of two groups of candidates taking the ABIM certifying examination for the first time between 1984 and 1988: graduates of Canadian medical schools and graduates of U.S. osteopathic medical schools. Numbers and types of residency programs attended are reported, and differences between program directors' rating of competence and examination performances are explored. Trends in numbers and performances are examined relative to those of two large baseline groups, USMGs and FMGs. Finally, the paper reports the joint effects of undergraduate training, type of residency program, and program director's rating of overall clinical competence on examination performance.

Method

This section describes the four groups studied and each measure on which they were compared.

Candidates

The candidates included in this study took the ABIM certifying examination in general internal medicine during the period 1984–1988. The study cohort in each year was limited to all first-time takers who had both finished their residencies and been evaluated by their program directors in the year of the examination. All the candidates had satisfactorily completed the equivalent of three years of training in a residency program accredited by the Accreditation Council for Graduate Medical Education (ACGME) or, if training had been completed in Canada, accredited either by The Royal College of Physicians and Surgeons of Canada or by the Professional Corporation of Physicians of Quebec. In addition, candidates with a Canadian undergraduate education must have attended a medical school accredited by the Committee for Accreditation of Canadian Medical Schools. Candidates from osteopathic medical schools must have graduated from colleges of osteopathic medicine accredited by the American Osteopathic Association.[13]

Types of Training

Residency programs were classified as (1) "top" U.S. programs (chosen by ballot of five members of the American Board of Internal Medicine, who were given a list of all accredited residencies), (2) university programs, and (3) non-university-affiliated programs. The second category consisted primarily of university and university-affiliated programs, although a small number of federal and military programs were included. The common feature of the final category was that all programs were independent of formal affiliations with universities. They were diverse, however, in that programs ranged from small hospitals training only a few residents per year to large programs that were well known and established.

Directors' Ratings

At the end of training, directors of internal medicine programs forwarded to the Board ratings of clinical competence for their residents who applied for the examination. Ratings were submitted for an overall evaluation of clinical competence and eight of its components: clinical judgment, medical knowledge, history-taking skill, physical examination skill, procedural skill, medical care, professional attitudes and behavior, and humanistic qualities. These evaluations were gathered on a nine-point scale where ratings of 1 to 3 were "unsatisfactory," 4 to 6 were "satisfactory," and 7 to 9 were "superior." Candidates who received overall ratings of 1 to 3 were not permitted to take the examination.

Examinations

The two-day written examinations were composed of patient management problems (in 1984–1986 only) and multiple-choice questions consisting of single-best-answer, matching (in 1984–1986 only), and multiple true-false items. A standardized score was developed for each reference group examinee for each item format. The reference group included all candidates who had graduated from U.S. or Canadian medical schools, were first-time takers of the certifying examination, and had recently completed three years of training in an approved residency program. Nonreference group members were given scores by locating their raw scores for each item format in the reference group distribution and assigning the associated standardized score. The standardized scores for each format were then weighted to form a composite which was restandardized to the reference group mean of 500 and a standard deviation of 100. This approach maintained the original rank ordering of candidates, as it placed all scores on the same scale. Additional details of the scoring process and estimates of reliability and validity can be found elsewhere.[6]

Results

In this section, tests of statistical significance are not provided. These data are based on the entire population of interest; consequently, there is no sampling error.

Numbers of Candidates

The number of Canadian graduates remained fairly stable, at about 100 candidates per year, except for a decrease to 72 in 1986. The number of osteopathic graduates nearly doubled over the five years, steadily increasing from 53 in 1984 to 102 in 1988. In comparison, the number of USMGs ranged from 3,435 to 3,659 and the

number of FMGs ranged from 859 to 1,061.

Program Types

During each year from 1984 to 1988, a little more than one-fifth of the USMGs attended a program ranked as a top U.S. program and a little less than one-fourth attended a non-university-affiliated residency program. Only 6% of the Canadian graduates attended the top U.S. programs over all five years, and similarly small percentages attended non-university-affiliated programs. By far, the most typical program was a university program. This is because about three-fourths of the Canadian graduates attended residency programs in Canada, all of which are university or university-affiliated programs. Osteopathic graduates were most likely to be in university programs and university-affiliated programs in 1984 and 1986, but overall, 50% were in non-university-affiliated programs. Finally, 60% of all the FMGs trained in non-university-affiliated programs.

Overall Clinical Competence

Over the five years, USMGs had an average rating of overall clinical competence of 6.7 (SD = 1.2). The Canadian graduates generally received ratings that fell just into the "superior" range of 7–9. Their lowest average was 6.9 (SD = 0.9) in 1987 and the highest was 7.2 (SD = 0.9), achieved in 1985. Average ratings for the osteopathic graduates were more variable, ranging from 6.3 (SD = 1.2) to 6.8 (SD = 1.1). The ratings for FMGs averaged 6.3 (SD = 1.2).

Components of Competence

Candidates were also rated by their program directors on eight components of clinical competence. Within each of the four subgroups of candidates, the five annual average ratings for each component were similar. Three observations were apparent: (1) for all the subgroups, the highest ratings were for humanistic qualities and professional attitudes and behaviors, while the lowest tended to be for history-taking, medical knowledge, physical examination skills, and procedural skills; (2) for each component, the Canadians received the highest average rating, followed by the USMGs, the osteopathic graduates, and the FMGs; and (3) the ordering of components within each subgroup was nearly the same for the USMGs, the Canadians, and the FMGs, whereas the osteopathic graduates were rated comparatively lower on medical knowledge but higher on procedural skills.

Examination Performance

Examination scores are shown in Table 1. Because the reference group was made up almost entirely of USMGs who were first-time takers of the examination, it follows that the average score for that group is always close to 500. Rankings of the four groups of candidates showed a consistent pattern within each of the five years. The highest scores were always achieved by the Canadian graduates, ranging from 524 (SD = 98) in 1984 to 570 (SD = 76) in 1987. The osteopathic graduates achieved average scores that were about one-half a standard deviation, or 50 points, lower than the average USMG scores and one-half a standard deviation higher than the average FMG scores.

The consistent pattern of differences observed among the four groups in types of residency programs attended, program directors' ratings, and average scores led the authors to perform analyses to explore how these factors, in combination with type of medical school attended, predicted examination scores. Data were combined over all five years and submitted to a multiple regression. The dependent variable was examination scores. The independent variables were (1) type of medical school, coded as three dummy (0/1) variables for Canadian, osteopathic, and foreign undergraduate medical education; (2) type of residency program; and (3) overall rating of clinical competence. Together, these variables accounted for 23.6% of the variability in the observed examination scores. The unique portions attributable to each variable were 4.5% of foreign medical schools, 0.2% to Canadian medical schools, 0.1% to osteopathic medical schools, 2.3% to type of residency, and 8.5% to program directors' ratings. The amounts of variation shared by osteopathic and Canadian medical schools were 0.1% with type of residency program and 0.2% with the overall rating of competence.

Table 1

Mean Scores of Four Groups of Medical School Graduates Who First Took the American Board of Internal Medicine Certifying Examination between 1984 and 1988*

	1984		1985		1986		1987		1988		Total	
Group	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
USMGs	499	100	499	101	500	100	498	100	498	100	499	100
Canadian	524	98	540	86	531	96	570	76	552	82	545	88
Osteopathic	440	93	452	104	446	108	452	116	465	108	453	107
FMGs	384	137	385	132	401	135	420	127	430	125	405	132

*The groups were United States medical school graduates (USMGs), Canadian medical school graduates, graduates of osteopathic schools in the United States, and graduates of foreign medical schools (FMGs).
Scores are on a scale standardized to a reference group's mean of 500 with a standard deviation of 100.

Conclusions

These findings are informative on several accounts. In addition to the fact that descriptive data regarding numbers of graduates and average examination performances are provided for two sizable groups of candidates that were previously unstudied, questions are raised about the importance of the graduates' different educational backgrounds. Undergraduate Canadian training and a residency in a Canadian university program are generally thought, at least by the ABIM, to be equivalent in quality to training in the United States. Although the separate effects of each are confounded in this study, data from this small subgroup of candidates support this assumption, and, in fact, suggest that the joint Canadian effect may even be better than a similar U.S. experience. The findings could, however, reflect a sampling issue. While over 200 residents seek internal medicine certification in Canada each year, only about half as many choose to sit for the ABIM examination.[14] The self-selection into the final step of the certifying examination process in this country may be more responsible for their outstanding performance than any specific facet of training.

The repeated finding that the osteopathic graduates, on average, did less well than did USMGs might also reflect a sampling issue, albeit in a less obvious way. Among the many reasons students attend osteopathic medical schools are because they are drawn to primary care and because of the unique philosophy of an osteopathic undergraduate education. The coupling of these two interests undoubtedly leads to diverse and important educational experiences. It might be, however, that additional areas of instruction that are also relevant to ABIM examination performance are less emphasized. This general condition of "less exam preparedness" may be exacerbated if the preference for general medicine steers osteopathic graduates away from the more tertiary-care-oriented university hospitals and toward the more community-oriented training programs. Unfortunately, the choice of attending a non-university-affiliated residency is, on average, associated with poorer certifying examination performance. At the risk of sounding too negative, however, it should be emphasized that the osteopathic graduates did well. Although not reported in the results section, their overall pass rate was 69% (compared with 84% for the reference group). This suggests that osteopathic graduates are an untapped reservoir of talented physicians. Stepped-up recruiting efforts directed at this group would seem a prudent strategy, given the internal medicine residency positions unfilled in the Match in recent years.[15]

This work was supported by the American Board of Internal Medicine but does not necessarily reflect its views or opinions.

References

1. Norcini, J. J., Grosso, L. J., Shea, J. A., and Webster, G. D. The Relationship Between Features of Residency Training and ABIM Certifying Examination Performance. *J. Gen. Intern. Med.* **2**(1987):330–336.
2. Norcini, J. J., Shea, J. A., Webster, G. D., and Benson, J. A., Jr. Predictors of the Performance of Foreign Medical Graduates on the 1982 Certifying Examination in Internal Medicine. *JAMA* **256**(1986): 3367–3370.
3. Shea, J. A., Norcini, J. J., Day, S. C., Webster, G. D., and Benson, J. A., Jr. Performance of U.S. Citizen Caribbean Medical School Graduates on the American Board of Internal Medicine Certifying Examinations, 1984 to 1987. *Teach. Learn. Med.* **1**(1989):10–15.
4. Benson, J. A., Jr., Meskauskas, J. A., and Grosso, L. J. Performance of U.S. Citizen–Foreign Medical Graduates on Certifying Examinations in Internal Medicine. *Am. J. Med.* **71**(1981):270–273.
5. Meskauskas, J. A., Benson, J. A., Jr., and Hopkins, E. Performance of Graduates of Foreign Medical Schools on the Examinations of the American Board of Internal Medicine. *N. Engl. J. Med.* **297**(1977): 808–810.
6. Norcini, J. J., Swanson, D. B., Grosso, L. J., and Webster, G. D. Reliability, Validity and Efficiency of Multiple Choice Question and Patient Management Problem Item Formats in Assessment of Clinical Competence. *Med. Educ.* **19**(1985):238–247.
7. Norcini, J. J., Swanson, D. B., Grosso, L. J., Shea, J. A., and Webster, G. D. A Comparison of Knowledge, Synthesis, and Clinical Judgment Multiple-Choice Questions in the Assessment of Physician Competence. *Eval. Health Prof.* **7**(1984): 485–500.
8. Webster, G. D., Shea, J. A., Norcini, J. J., Grosso, L. J., and Swanson, D. B. Strategies in Comparison of Methods for Scoring Patient Management Problems. *Eval. Health Prof.* **11**(1988):231–248.
9. Norcini, J. J., Fletcher, S. W., Quimby, B. B., and Shea, J. A. Performance of Women Candidates on the American Board of Internal Medicine Certifying Examinations, 1973–1982. *Ann. Intern. Med.* **102**(1985): 115–118.
10. Appendix III, Medical Schools in Canada. *JAMA* **260**(1988):1157–1161.
11. Statistical Tables on the Osteopathic Profession. In *1988–89 Yearbook and Directory of Osteopathic Physicians.* Chicago, Illinois: American Osteopathic Association, 1988.
12. Crowley, A. E., and Etzel, S. I. Section III: Graduate Medical Education in the United States. *JAMA* **260**(1988):1093–1101.
13. *Policies and Procedures.* Philadelphia, Pennsylvania: American Board of Internal Medicine, July 1988.
14. *Annual Report and Performance Handbook: 1987.* Ottawa, Canada: Royal College of Physicians and Surgeons, 1988.
15. Aday, L. A., et al. National Study of Internal Medicine Manpower XII. The Future of Graduate Medical Education in Internal Medicine: What Do Program Directors Predict? *Arch. Intern. Med.* **148**(1988): 1509–1514.